HAL
open science

Calibración de cámaras de tiempo de vuelo: Ajuste adaptativo del tiempo de integración y análisis de la frecuencia de modulación

P. Gil, T. Kisler, G.J. García, C.A. Jara, J.A. Corrales

To cite this version:

P. Gil, T. Kisler, G.J. García, C.A. Jara, J.A. Corrales. Calibración de cámaras de tiempo de vuelo: Ajuste adaptativo del tiempo de integración y análisis de la frecuencia de modulación. Revista Iberoamericana de Automática e Informática Industrial RIAI, 2013, 10 (4), pp.453 - 464. 10.1016/j.riai.2013.08.002 . hal-01680652

HAL Id: hal-01680652

https://uca.hal.science/hal-01680652v1

Submitted on 15 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ScienceDirect
Disponible en www.sciencedirect.com
Revista Iberoamericana de Automática e Informática industrial 10 (2013) 453–464
www.elsevier.es/RIAI

Calibración de cámaras de tiempo de vuelo: Ajuste adaptativo del tiempo de integración y análisis de la frecuencia de modulación

P. Gil [a,*], T. Kisler[b], G.J. García[a], C.A. Jara[a], J.A. Corrales[a]

[a] *Departamento de Física, Ingeniería de Sistemas y Teoría de la Señal, Universidad de Alicante, Crta/ San Vicente del Raspeig s/n, 03690 Alicante, España.*
[b] *Universidad Técnica de Múnich, Múnich, Alemania.*

Resumen

La percepción de profundidad se hace imprescindible en muchas tareas de manipulación, control visual y navegación de robots. Las cámaras de tiempo de vuelo (ToF: Time of Flight) generan imágenes de rango que proporcionan medidas de profundidad en tiempo real. No obstante, el parámetro distancia que calculan estas cámaras es fuertemente dependiente del tiempo de integración que se configura en el sensor y de la frecuencia de modulación empleada por el sistema de iluminación que integran. En este artículo, se presenta una metodología para el ajuste adaptativo del tiempo de integración y un análisis experimental del comportamiento de una cámara ToF cuando se modifica la frecuencia de modulación. Este método ha sido probado con éxito en algoritmos de control visual con arquitectura 'eye-in-hand' donde el sistema sensorial está compuesto por una cámara ToF. Además, la misma metodología puede ser aplicada en otros escenarios de trabajo. *Copyright © 2013 CEA. Publicado por Elsevier España, S.L. Todos los derechos reservados.*

Palabras Clave:
Tiempo de vuelo, calibración, imagen de rango, percepción robótica, cámaras 3D.

1. Introducción

Tradicionalmente, en tareas de control visual y manipulación robótica o en la navegación de robots autónomos, se ha empleado medidas de radar, sonares, sensores láser, dos o más cámaras CCD o CMOS formando sistemas binoculares, trinoculares, etc. Sin embargo, algunos de estos sistemas basados en cámara tienen una fuerte limitación cuando los objetos de la escena empleados como objetivo carecen de una textura homogénea o bien la escena tiene unas condiciones extremas de baja iluminación. En particular, el uso de sensores CCD o CMOS ha sido empleado ampliamente en arquitecturas de control visual basados en imagen (Chaumette *et al.*, 2006). No obstante, cuando se ha empleado este tipo de sensores en control visual basado en posición, se ha requerido de una estimación de la distancia entre cámara y objeto. En los últimos años, la aparición de la tecnología ToF (ToF: Time of Flight) ha permitido estimar la distancia entre objeto y cámara a partir de imágenes de profundidad de la escena, por lo que solventa algunos de los problemas comentados (Pomares *et al.*, 2010). Este tipo de sensores comienza a emplearse como elemento indispensable en sistemas hombre-máquina (Schwarz *et al.*, 2010) (Shotton *et al.*, 2011), en la industria del entretenimiento (Khoshelham *et al.*, 2011), como parte de sistemas de sensorizado en aplicaciones de realidad aumentada (Kolb *et al.*, 2010), o como ya se ha comentado, como parte del sistema de percepción de sistemas robóticos (Hussman *et al.*, 2009, 2010). A veces se emplea como parte única del sistema sensorial (Kuehnle *el al.*, 2008) y otras muchas en combinación con otros sensores (Zhu *et al.*, 2008, 2011). En aquellos casos, en los que la cámara ToF forma parte del sistema de percepción de un robot que se mueve, la escena varía y la distancia de los objetos presentes cambia continuamente. En la mayoría de estos casos, se requiere adaptar el tiempo de integración (May *et al.*, 2006) (Wiedemann *et al.*, 2008) (Gil *et al.*, 2010) y/o la frecuencia de modulación de la cámara para mejorar la calidad de la imagen y disminuir el error en el cálculo de las distancias (Mufti *et al.*, 2010) (Pattison *et al*. 2010).

Por un lado, la principal ventaja de una cámara ToF es la extracción de profundidad para determinar características geométricas y planos de la escena de trabajo. Además, una correcta calibración de ésta, permite transformar la imagen de profundidades en un conjunto de puntos 3D del entorno (Fuchs *et al.*, 2008), proporcionando así información espacial tridimensional. Por otro lado, el principal inconveniente son las fluctuaciones en la precisión con la que se mide la distancia al objetivo, principalmente debido a las interferencias causadas por factores externos, como por ejemplo la luz solar, la orientación, la reflectividad de la superficie de los objetos de la escena y la distancia de trabajo entre cámara y objetivo. En la literatura, varios son los autores que han hecho estudios que modelan y

* Autor en correspondencia.
Correos electrónicos: pablo.gil@ua.es (P. Gil)
URL: www.aurova.ua.es/pgil (P. Gil)

© 2013 CEA. Publicado por Elsevier España, S.L. Todos los derechos reservados
http://dx.doi.org/10.1016/j.riai.2013.08.002

estiman el error en cámaras ToF (Rapp *et al.*, 2008) (Pattison *et al.*, 2010).

Actualmente, los sensores ToF han alcanzado un gran impacto en diferentes áreas de investigación de la visión por computador y la percepción, desde el reconocimiento de objetos (Lai *et al.*, 2011), la prevención de colisiones y/o la reconstrucción de escenas y/u objetos (Kakiuchi *et al.*, 2010). En todos ellos se requiere medir distancias con una buena precisión. De ahí, que sean variados y numerosos los estudios cuyo objetivo es la obtención de métodos de calibración que permitieran medir distancias a objetos con la mayor exactitud posible (Weyer *et al.*, 2008) (García *et al.*, 2011). Para ello, en estos últimos cuatro años especialmente, se ha buscado corregir los errores sistemáticos que se producen en este tipo de cámaras como consecuencia de una mala calibración o una ausencia de ella. Algunos de estos estudios han buscado calibrar una cámara ToF fusionando la información que ésta proporciona con la de otro tipo de sensores, tales como una cámara CCD/CMOS (Kim *et al.*, 2008), con múltiples cámaras (Schiller *et al.*, 2008), o con cámaras color RGB (Herrera *et al.*, 2011) (Shahbazi *et al.*, 2011). Esto se ha llevado a cabo tanto en sensores ToF de uso industrial (PMD, CSEM, etc.) como de bajo coste y uso doméstico como Kinect de Microsoft (Smisek *et al.*, 2011). Otros investigadores han presentado nuevas estrategias que buscan obtener métodos de autocalibración sin requerir de sensores adicionales y fusión de datos (Lichti *et al.*, 2008, 2009, 2011) (Lindner *et al.*, 2008, 2010).

Además, en aquellas aplicaciones dónde se requiere seguir un objeto o reconstruir y modelar un entorno, surgen otras dificultades. Cuando en la escena hay movimiento, la cámara ToF genera un video formado por los mapas de profundidades de las imágenes de la secuencia capturada. Esta secuencia de imágenes requiere de un proceso de correspondencias que permita establecer qué región de una imagen se corresponde con la misma región en la otra imagen. En el caso de reconstrucción de entornos, existen en la bibliografía diversos métodos que permiten obtener una representación 3D. Ejemplo de ellos son las técnicas basadas en el algoritmos ICP (ICP: Iterative Closest Point; puntos más cercanos iterativos) (Zinber *et al.*, 2003) que consiste en encontrar las relaciones entre las nubes de puntos que definen regiones de objetos. En ocasiones, estos algoritmos se combinan con técnicas basadas en SLAM (SLAM: Simultaneous Locatization And Mapping) donde se requiere localizar los objetos presentes en la escena, localizar el robot y evitar los obstáculos que aparecen en la trayectoria de éste mediante la fusión de información en el tiempo (May *et al.*, 2009a,b).

En este artículo se presenta un método para el autoajuste adaptativo del tiempo de integración de cámaras ToF. El método propuesto se basa en el análisis de los errores obtenidos en experimentos de posicionamiento de un brazo robot. La validez de la metodología presentada se muestra aplicándola a tareas de control visual que tiene particularidades muy concretas como rangos de trabajo variable e inferiores a 1 metro. El autoajuste adaptativo del tiempo de integración y la elección adecuada de la frecuencia de modulación, permiten integrar sensores ToF en aplicaciones de guiado de robots empleando técnicas de control de posición que requieran precisión, no sólo en posiciones en un plano, x e y sino también precisión en el posicionamiento en profundidad, z. Además, la metodología presentada es válida para ajustar estos dos parámetros, tiempo de integración y frecuencia de modulación, a otros entornos de trabajo.

Este artículo se ha organizado de la siguiente manera. En la sección 2 se comentan los principios básicos de funcionamiento de la tecnología ToF y una comparativa de las características de las cámaras ToF existentes en el mercado. En las siguientes secciones se muestran y discuten resultados experimentales que indican cómo influyen los parámetros configurables de la cámara (tiempo de integración y frecuencia de modulación) en el cálculo de la distancia entre cámara y objetivo. Así, en la sección 3, se muestra el comportamiento de este tipo de cámaras cuando se trabaja a distancias inferiores a un metro y se presenta una metodología para modelar el tiempo de integración más adecuado y que éste pueda adaptarse según las distancias de trabajo. De modo similar, en la sección 4, se explica el comportamiento de la cámara ToF en función de la frecuencia de modulación. Finalmente, en la sección 5 se presentan las conclusiones de estos trabajos.

2. Descripción de una cámara ToF

2.1. Principios físicos de funcionamiento

Las cámaras ToF emplean luz modulada en amplitud. Estas cámaras montan cerca de la óptica, matrices de diodos de luz que emiten señales con longitud de onda cercana al infrarrojo (Figura 1). Estas fuentes de luz son de tipo NIR (NIR: Near Infrared; infrarrojo cercano). Las cámaras basadas en el principio de tiempo de vuelo miden el tiempo que necesitan los haces en propagarse como luz emitida hasta la superficie del objeto y en regresar en forma de luz reflejada. En definitiva, el sensor que montan estas cámaras estima el desplazamiento en fase mediante correlación entre la señal reflejada y la señal empleada como referencia. A este fenómeno se le conoce como Interferometría de desplazamiento en fase (PSI: Phase Shifting Interferometry). Un estudio de interferometría en cámaras ToF es expuesto en (Frank *et al.*, 2009).

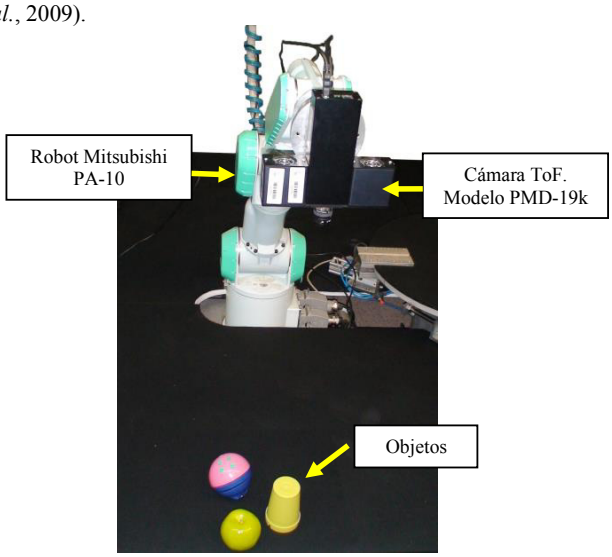

Figura 1: Sistema robótico PA-10, cámara de tipo ToF (Modelo PMD-19K) y entorno de trabajo.

Matemáticamente, conocida la velocidad de la luz, c, la frecuencia de modulación, f_{mod}, la correlación entre cuatro señales desplazadas fases de 90º, denotadas por $r_0(0º)$, $r_1(90º)$, $r_2(180º)$, $r_3(270º)$ y tomadas con un periodo de $T=1/f_{mod}$ es posible calcular

el retardo en fase, ϕ_i, la amplitud, a_i, y la distancia , d_i, en cada píxel i de los n píxeles que forman la imagen bidimensional de la siguiente manera:

$$\phi_i = arctan\left(\frac{r_1 - r_3}{r_0 - r_2}\right) \quad (1)$$

$$a_i = \frac{\sqrt{(r_1 - r_3)^2 + (r_0 - r_2)^2}}{2} \quad (2)$$

$$d_i = \frac{c \cdot \phi_i}{4\pi \cdot f_{mod}} \quad (3)$$

2.2. Imágenes ToF

En la Figura 2, se muestra una imagen de pseudo-color generada a partir de la fusión de la información en escala de grises adquirida con una cámara ToF (Figura 1) y la información de distancia que proporciona el mapa de profundidades, es decir la distancia d_i calculada para cada píxel de acuerdo con (3). En la imagen generada se ha empleado el factor de amplitud, a_i, calculado con (2), para marcar cómo de fiables y representativos son los valores de profundidad estimados para los distintos píxeles, i, de la imagen. Así, las regiones que se han rellenado con un pseudo-color rosa son las regiones de mayor amplitud y por lo tanto más representativas del objeto. En la Figura 2, es posible observar que una cámara ToF se puede emplear para mejorar los algoritmos de segmentación entre objetos y fondo. El algoritmo empleado para segmentar objetos de fondo (Kisler et al., 2011) se ha comentado en el apéndice de este artículo.

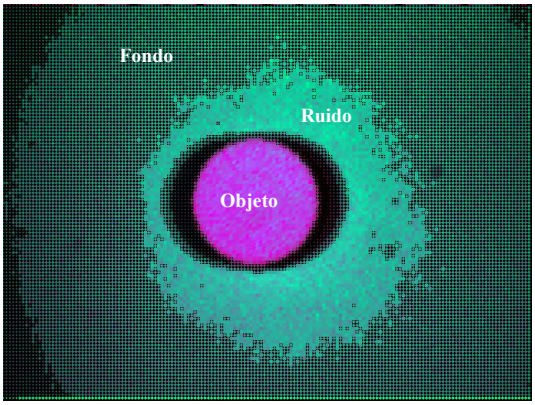

Figura 2: Imagen adquirida con una cámara ToF (Modelo PMD-19K).

2.3. Parámetros activos: tiempo de integración y frecuencia de modulación

Como se ha comentado, las cámaras ToF emiten luz y calculan una estimación de la distancia al objetivo, d_i, basándose en el tiempo que tarda la luz emitida en ser reflejada y posteriormente recibida en el sensor de la cámara. Las cámaras ToF tienen dos parámetros que determinan la precisión y robustez del cálculo de la distancia. El primero de ellos, el tiempo de integración, τ, que define el tiempo que requiere el sensor de la cámara para recibir la señal de luz reflejada por la superficie de los objetos que se han establecido como objetivos en la escena. Esto significa que si el tiempo de integración es pequeño, la distancia a los objetos muy alejados no podrá ser medida, correctamente, puesto que la señal de luz no habrá tenido tiempo de retornar y excitar el sensor de la cámara. Además, por otro lado, si el tiempo de integración es grande, la cámara reduce su velocidad de adquisición y los mapas de profundidades pueden reflejar valores de distancia incorrectos, generalmente sobresaturados (Wiedemann *et al.*, 2008) (Gil *et al.* 2010). El segundo parámetro ajustable es la frecuencia de modulación, f_{mod}, de la luz emitida. La variación de este último parámetro cambia la apariencia física de la forma coseno que modela matemáticamente esa señal de luz (Figura 3).

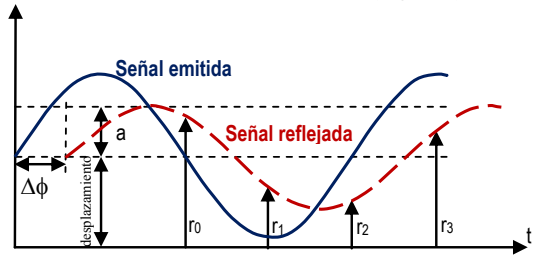

Figura 3: Señal emitida y reflejada por la cámara ToF.

Este parámetro determina la máxima distancia al objetivo que la cámara es capaz de estimar. Por ejemplo, el fabricante de la cámara PMD-19k emplea por defecto una frecuencia de modulación de 20Mhz, con la que se limita la distancia máxima a estimar en 7.5m. Este valor se conoce con el sobrenombre de 'rango de no ambigüedad' (Schaller *et al.*, 2011). Este rango de no ambigüedad (Tabla 1) se puede calcular como:

$$r_{na} = \frac{c}{2 \cdot f_{mod}} \quad (4)$$

En las siguientes secciones, se presentan varios experimentos que analizan el comportamiento del tiempo de integración y cómo influye éste en el cálculo de la distancia entre cámara y objetos de interés en la escena. De dichos experimentos se puede deducir que una buena estimación de un tiempo de integración adecuado es un factor crítico a la hora de calcular la distancia, d_i, de la ecuación (3). Además, se presentan una serie de experimentos con el objetivo de validar o desaprobar la recomendación del fabricante de emplear una frecuencia de modulación por defecto de 20Mhz (Tabla 2).

2.4. Prestaciones

Actualmente, es habitual que muchos de los sistemas de percepción integren cámaras ToF. Entre los fabricantes de este tipo de cámaras destacan PMDTechnologies y CSEM (Foix *et al.* 2011). En las tablas 1 y 2 se muestran las características y especificaciones de las principales cámaras de estos fabricantes. Unos modelos difieren de otros en la resolución de imagen dependiente del tamaño de los sensores, la velocidad de adquisición de imágenes a máxima resolución, la potencia de la fuente de iluminación y la longitud de onda de la señal modulada empleada para aplicar el principio de ToF (Figura 3). Estas especificaciones y la calidad constructiva de la cámara determinan algunas de las prestaciones de éstas.

Notar que en los experimentos que se presentan en este artículo se ha empleado una cámara PMD-19K que proporciona una imagen segmentada de distancia (Figura 2). El proceso de segmentación se muestra en el apéndice.

Tabla 1: Prestaciones de cámaras ToF en el mercado

Fabricante/Modelo	Rango no Ambigüedad	Resolución de distancia	Error d_i
PMD-19K	7.5m	3%	6mm
PMD-CamCube	0.3-7m	2%	3mm
PMD-CamBoard	2m	1%	1mm
CSEM-SR3000	7.5m	1%	1mm
MESA-SR4000	0.3-5m	1%	1mm

Tabla 2: Especificaciones de cámaras ToF en el mercado

Fabricante/ Modelo	Iluminación	Velocidad/ Resolución	f_{mod}
PMD-19K	3W-870nm	15fps/160x120	20Mhz
PMD-CamCube	3W-870nm	40fps/200x200	20Mhz
PMD-CamBoard	1W-850nm	90fps/160x120	30Mhz
CSEM-SR3000	1W-850nm	25fps/176x144	20Mhz
MESA-SR4000	1W-850nm	50fps/176x144	30Mhz

Conviene resaltar que algunos de los inconvenientes de las cámaras ToF (Distante *et al*, 2010), desde el punto de vista de los sensores que emplean respecto a otros tipos de cámaras como las CCD o CMOS, radican en: su baja resolución y baja velocidad de adquisición, distorsión en el cálculo de la distancia, d_i, valores de distancia erróneos en los contornos de los objetos e interferencias entre la iluminación del entorno y la de la cámara u otras cámaras. Además, el coste de los sensores ToF por lo general es bastante más elevado. Pero también, proporcionan ciertas ventajas entre las que cabe destacar: trabajan en ausencia de iluminación ambiental, no les afectan las oclusiones parciales de los objetos debido a sombras u a otros objetos, no les influyen las propiedades de apariencia de los objetos como las formas, aunque sí puede afectar el material y la cromaticidad de éstos.

3. Ajuste del Tiempo de Integración

3.1. Análisis de comportamiento de cámara ToF en función del tiempo de integración

El tiempo de integración, τ, es el parámetro activo más importante de una cámara ToF. La estabilidad de las medidas obtenidas, tales como las intensidades de los píxeles o la distancia asociadas a éstos, dependen del tiempo de integración. En algunas cámaras ToF, como por ejemplo PMD19K, este parámetro se debe ajustar previamente de modo manual (calibración fuera de línea) y en otras cámaras como las SR4000 este parámetro es calculado automáticamente (calibración en línea, on-line) por el propio software de la cámara. No obstante, τ debe ajustarse en función del tipo de escena y de la distancia entre objetos que aparecen en el entorno, ya sea manual o automáticamente. Aunque algunas cámaras, como la SR4000, auto-ajustan este parámetro no siempre el ajuste determinado por el fabricante es el mejor posible. Esto es así, porque las condiciones del entorno de trabajo y de los objetos de trabajo pueden ser muy diversas.

Consecuentemente, cuando una cámara ToF se monta en el efector final de un robot o sobre cualquier plataforma móvil, τ tiene que ser ajustado en un proceso en línea. Esto es así, porque hay movimiento en la escena y las distancias de trabajo varían. Esto implica que la percepción que se tiene del entorno sea distinta, cambian posiciones y escalados al variar el punto de vista de la cámara ToF con el movimiento. Por lo tanto, se hace necesario determinar una metodología algorítmica para ajustar automáticamente este parámetro (on-line) minimizando el error en el cómputo de la distancia, d_i. y adecuándolo a las variaciones de percepción. Así, si τ se estima adecuadamente, la adquisición de imágenes se realizará en las mejores condiciones posibles, y esto repercutirá en la calidad de las imágenes de intensidades adquiridas, aunque varíe el punto de vista de la cámara debido al movimiento. En la Figura 4, se muestran varias imágenes de un mismo objeto con distintos valores de τ_i sin variar las condiciones de la escena (fondo, distancias de trabajo, iluminación, punto de vista, etc.). El objeto se ha colocado sobre un fondo negro que produce baja reflectividad en los bordes del objeto. Para controlar la distancia, la cámara se ha montado sobre el efector de un robot manipulador (Figura 1) y para controlar la luz ambiental, se ha empleado un regulador de corriente que permite controlar las fuentes de luz continua empleadas en el entorno de trabajo. De este modo, se consigue limitar las interferencias de las fuentes de luz NIR de la cámara ToF.

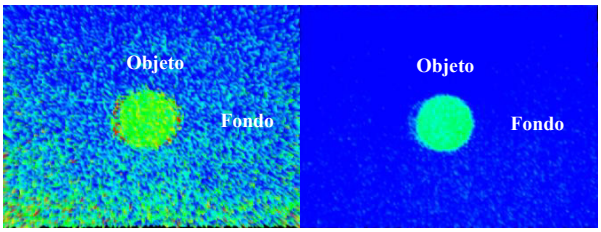

Figura 4: Imágenes capturadas con distintos valores para el tiempo de integración (10ms y 30ms).

Para determinar el rango de valores a los que interesa ajustar el tiempo de integración, τ, para que la cámara ToF adquiera imágenes de distancias, I, de buena calidad, dónde se reduzca el número de píxeles con errores de medida importantes, se ha procedido de la siguiente manera:

Algoritmo 1: Estimar Comportamiento ToF

1. Para k=1 hasta N ∀ N posiciones del robot
   a. Mover el robot a la posición k / k sitúa el robot a una distancia d ∈ [200,1000]mm.
   b. Para τ=0 hasta 80ms.
      i. Adquirir mapa de distancias d_i ∀ pixel-i ϵ I
      ii. Calcular la distancia media d_m (Figura 5)
      iii. Adquirir mapa de intensidades l_i ∀ pixel-i ϵ I
      iv. Calcular la intensidad media l_m (Figura 6)

En las Figuras 5 y 6 se observa como si τ es pequeño, la distancia y la intensidad que se computan es inestable o poco fiable, respectivamente. Por otro lado, si τ es grande se produce, en ocasiones, un fenómeno de sobresaturación. Normalmente, se puede afirmar que este fenómeno aparece cuando la distancia entre cámara y objetivo es muy pequeña. Empíricamente se ha calculado que ese valor de umbral que limita la zona de

sobresaturación esta cercano al límite inferior del rango de no ambigüedad. En el ejemplo de la Figura 5, líneas roja y morada, este valor es de 0.4m. Sin embargo, cuando la cámara se posiciona lejos del objetivo las distancias computadas por la cámara son próximas a las reales que determinan la posición de la cámara. Algo similar se puede observar en la Figura 6.

La distancia media, d_m, cuya evolución se muestra en la Figura 5, se ha calculado como la media aritmética de la distancia asociada a cada píxel, d_i, del mapa de distancias adquirido a partir de la medida computada por el sensor ToF de acuerdo a (3). Es decir:

$$d_m = \sum_{i=1}^{n \cdot m} d_i / n \cdot m \tag{5}$$

De modo similar, a partir de los valores de intensidad de cada píxel en la imagen se computa la intensidad media, i_m. Su evolución en función del tiempo de integración empleado por el sensor ToF se muestra en la Figura 6.

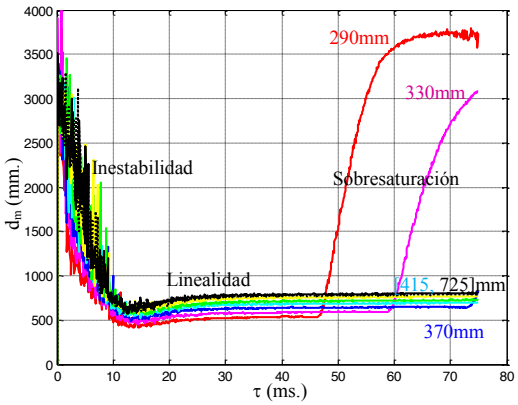

Figura 5: Evolución del cálculo de d_m en la ejecución del algoritmo 1, para posiciones del robot entre [200,1000]mm.

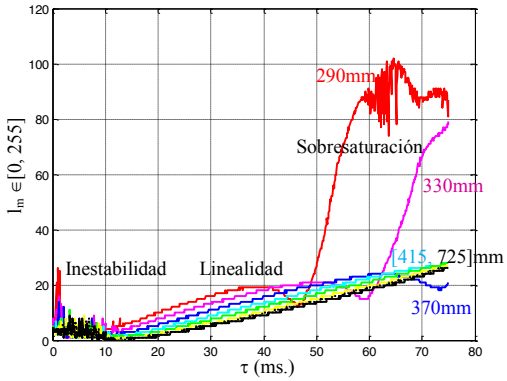

Figura 6: Evolución del cálculo de I_m en la ejecución del algoritmo 1, para posiciones del robot entre [200,1000]mm.

Por lo tanto, se llega a la conclusión que es importante determinar un umbral inferior y uno superior que limiten la zona de uso eficaz del valor de tiempo de integración, τ, sin que se produzcan errores importantes en el cálculo de la distancia.

En la metodología mostrada en el algoritmo 1 se emplean valores medios de distancia e intensidad para conseguir independencia del número de objetos y localización de estos en la imagen (posición y escalado). Los valores medios son más robustos y estables ante errores de medida locales asociados a píxeles. Los valores medios se pueden computar a partir de toda la imagen o a partir de la región segmentada para los objetos presentes en ella (ver Apéndice A). Así, en la metodología propuesta siempre que la escena de trabajo sea desconocida y difícilmente segmentable, porque presenta muchos objetos en planos distintos y con texturas dispares, entonces se utilizan valores medios globales de toda la imagen. Si por el contrario, la escena no es excesivamente compleja, caso de guiado con control visual donde se parte de objetos conocidos que se emplean como marcadores de guiado, entonces los valores medios son locales. Éstos se calculan a partir de las regiones segmentadas que representan las áreas de los objetos presentes en la escena. Para segmentar las regiones de los objetos se puede emplear el algoritmo del Apéndice A.

3.2. Cálculo del tiempo de integración

En la Figura 5, se puede comprobar que la zona útil de tiempos de integración viene dada por la zona de linealidad. Realizando un estudio similar al mostrado en las Figuras 5 y 6 pero en este caso para el parámetro amplitud computado por el sensor ToF como se indica en (2), se obtiene la amplitud media, a_m, mediante:

$$a_m = \sum_{i=1}^{n \cdot m} a_i / n \cdot m \tag{6}$$

Por lo general, a mayor valor de amplitud para un píxel más fiable es el valor distancia calculado para éste. La amplitud es, por lo tanto, un parámetro de bondad para cada posible tiempo de integración, τ. Si se analiza la información que proporciona la amplitud en función del τ empleado (Figura 7) y se compara con la proporcionada por la figura 5, se puede observar cómo los valores útiles de tiempo de integración corresponden a aquellos que se encuentran entre el mínimo y máximo de la señal gaussiana que representa cada iteración k del Algoritmo 1. La señal gaussiana representa la amplitud media calculada con (6) para cada posición k del robot. Cada posición del robot, determina una distancia real entre cámara y objetivo.

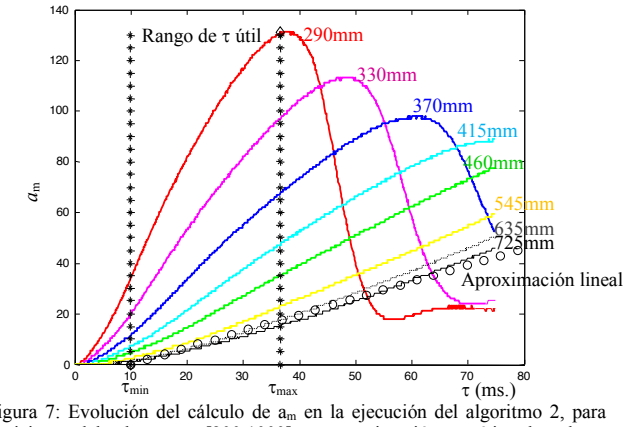

Figura 7: Evolución del cálculo de a_m en la ejecución del algoritmo 2, para posiciones del robot entre [200,1000]mm. y estimación empírica de valores útiles de tiempo de integración.

De acuerdo, al análisis de las medidas que se muestran en la figura 7, se observa que cuando la distancia real entre cámara y objetivo es pequeña, la señal de amplitud media a_m tiene una forma gaussiana, donde cada curva gaussiana representa una distribución frecuencial. Además, se comprueba que la curva crece más rápidamente y es fácilmente perceptible su máximo absoluto. Sin embargo, cuando la distancia entre cámara y objetivo es grande no hay un máximo absoluto.

Este valor de máximo define un umbral para delimitar el tiempo máximo de integración, τ_{max}, que se considera adecuado para la adquisición de imágenes. Por encima de este valor pueden producirse fenómenos de sobresaturación que impiden un correcto cómputo de la distancia de cada píxel. Además, del tiempo máximo de integración, es importante fijar un tiempo de integración mínimo, τ_{min} que delimite cual es el valor más pequeño que se recomienda usar para la adquisición de imágenes. El tiempo de integración deseado, τ_d, debe estar dentro del rango, $[\tau_{min}, \tau_{max}]$. De este modo, se garantiza un cálculo adecuado de la distancia e intensidad en cada píxel, y puede ser calculado añadiendo el cálculo de a_m y dos nuevos pasos al algoritmo 1. En tal caso, el nuevo algoritmo queda del siguiente modo:

Algoritmo 2: Estimar Rango útil de τ

1. Para k=1 hasta N /∀ N posiciones del robot
   a. Mover el robot a la posición k / k sitúa el robot a una distancia d ∈ [200,1000]mm.
   b. Para τ=0 hasta 80ms.
      i. Adquirir mapa de amplitudes a_i ∀ pixel-i ϵ I
      ii. Calcular la amplitud media a_m (Figura 7)
2. Calcular τ_{min} tal que $a(\tau_{min})=\min\{a_m\}$ cuando k=N. Es decir cuando el robot está a la máxima distancia de trabajo.
3. Calcular τ_{max} tal que $a(\tau_{max})=\max\{a_m\}$ si el objeto está cerca del objetivo o $a(\tau_{max})$=cuartil_superior$\{a_m\}$ si está lejos del objetivo.
4. Escoger $\tau_d=\tau_{max}$

Finalmente, se ha evaluado el método de ajuste automático del tiempo de integración en una tarea de control visual. Para ello se ha empleado la adquisición de imágenes en una secuencia en movimiento, en la que en cada iteración, la distancia entre cámara ToF y objetivo varía. Así, es importante notar, que en la metodología mostrada en el algoritmo 2 es posible modificar el rango de distancia para posicionar el robot de acuerdo a las especificaciones de uso en las que se vaya a emplear la cámara ToF. Esta modificación no influye en la metodología de estimación de τ.

El movimiento simula una trayectoria en la que un robot manipulador controla su posicionamiento en función de las características visuales que proporciona la cámara ToF (Figura 8). El esquema para llevar a cabo este control en posición viene determinado por una arquitectura de control visual 'eye in hand' en la que la cámara ToF se monta sobre el extremo del robot como se indica en la Figura 1.

En este experimento el extremo del robot se ha movido con una velocidad constante de 0.01m/s. Y se ha ejecutado una trayectoria de acercamiento, en la que la posición inicial del extremo del robot es de 830mm y la posición final es de 240mm (Figura 8).

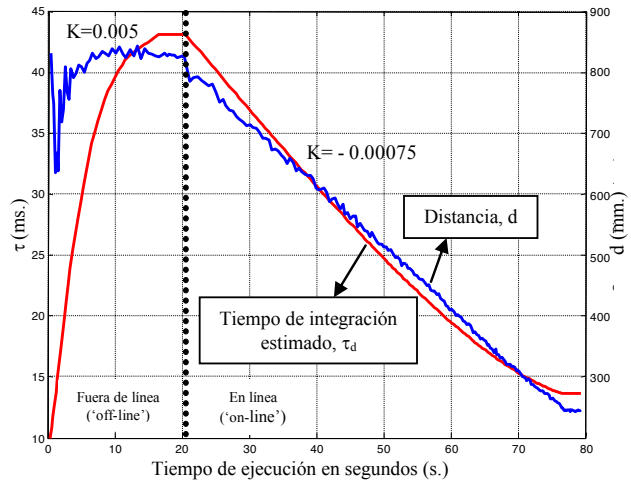

Figura 8: Error en la adaptación del tiempo de integración en función de la distancia entre cámara y objetivo.

El tiempo de integración para la posición inicial, se ha calculado a partir del algoritmo 2. Una vez calculado este tiempo de integración, este se ha ido actualizando a partir de la siguiente ley de control:

$$(\tau_d)_{k+1} = (1 + K \cdot e) \cdot (\tau_d)_k \quad (7)$$

donde el valor de e, representa la desviación existente entre el valor de amplitud medio deseado para la imagen que se capturaría escogiendo el tiempo de integración ideal para esa posición y el que realmente se ha estimado con el algoritmo 2. En este caso, para el caso en el que se parte de una posición alejada entre cámara y objeto:

$$e = max\{a_m(\tau_k)\} - a_m(\tau_k) \quad (8)$$

Así, por un lado, en la Figura 8 se muestra la estimación del tiempo de integración para la posición inicial, en un proceso de pre-ajuste o calibración offline, a partir del algoritmo 2. Por otra parte, la misma figura muestra cómo el tiempo de integración se actualiza, adaptándose a la distancia de trabajo cuando el robot se mueve en la trayectoria de acercamiento definida. En este experimento se ha escogido una K=0.00075 para adecuar la velocidad de actualización del parámetro, $(\tau_d)_k$, a la velocidad de movimiento del extremo del robot.

Es más, a partir de los análisis llevados a cabo y comentados en este apartado es factible construir un modelo matemático que permita asociar un tiempo de integración subóptimo, τ_d^*, para cada distancia (Gil *et al.*, 2010). Así asumiendo que conocemos el tiempo de integración deseado, para varias posiciones, se puede estimar la variación del tiempo de integración en función de la distancia entre cámara y objetivo mediante aproximación polinómica. Con la finalidad de construir una tabla de valores que permita actualizar, rápidamente, el tiempo de integración en función de la distancia de trabajo, sin necesidad de precalibrar la cámara para fijar un parámetro inicial de tiempo de integración de acuerdo al algoritmo 2.

$$\tau_d^* = 0.57d^5 + 1.14d^4 - 1.53d^3 + 0.25d^2 + 11.88d + 20.6 \quad (9)$$

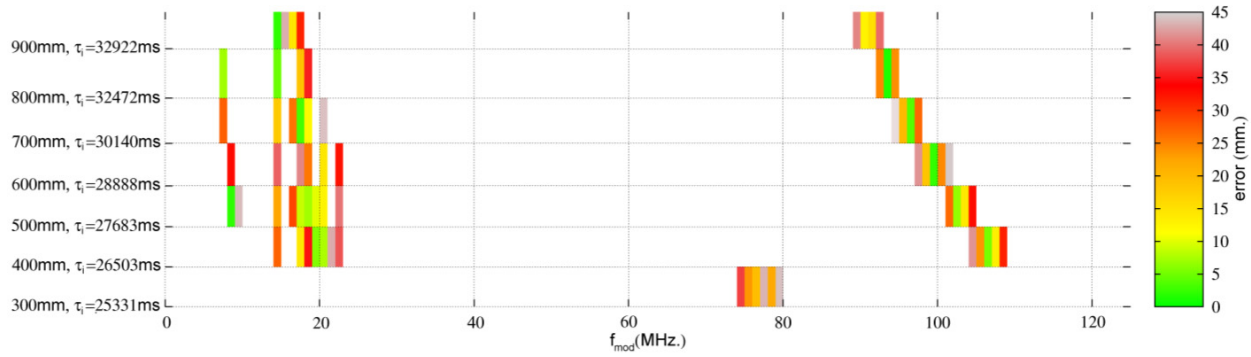

Figura 9: Evaluación del error en la distancia estimada entre observador y objetivo cuando se usa un tiempo de integración adaptativo, τ_k, y se varía f_{mod}.

3.3. *Análisis del error*

En esta sección se muestra un análisis del comportamiento del error en el proceso de localización de objetos, presentado en la sección anterior. Para ello se ha empleado el tiempo de integración adaptativo, calculado en (9). La figura 9 sólo representa los errores de distancia computados para los rangos de frecuencia de modulación entre 17-20Mhz y 90-110Mhz. Es en estos rangos de frecuencia de modulación donde los errores se mantienen contenidos por debajo de los 45mm. En el resto de frecuencias los errores se incrementan de forma muy elevada y no se ha considerado su representación en la figura.

Así, en la Figura 9, se puede observar cómo para una frecuencia de modulación, f_{mod}, en torno a los 20Mhz (valor recomendado por el fabricante) y en función del tiempo de integración, τ_k, empleado para distintas distancias de trabajo, el error siempre se mantiene dentro de un rango de valores aceptable. Aunque no es posible definir una dependencia entre la frecuencia y el tiempo de integración, no obstante sí hay dependencia entre el tiempo de integración y la distancia como se demostró en la sección anterior. Como se observa, la precisión alcanzada varía significativamente para diferentes frecuencias de modulación como muestran los colores de la gráfica que definen el comportamiento del error, los cuales varían con aleatoriedad. Además, en torno a los 100Mhz de frecuencia de modulación, se puede muestra como el comportamiento del error se hace más estable y más predecible. El error se hace inferior a 5mm (color verde), siempre que la frecuencia de modulación empleada se haga crecer inversamente proporcional al tiempo de integración empleado. De aquí se deduce que es interesante trabajar con frecuencias de modulación altas cuando la cámara con la que se observa la escena se mantiene a distancias de trabajo inferiores a 1 metro, ya que se puede extraer una relación de proporcionalidad entre tiempo de integración y frecuencia de modulación. Distancias en las que suele trabajar en aplicaciones de control visual (Pomares *et al.*, 2010), aunque el error no esté próximo a la máxima precisión alcanzada definida por el fabricante de la cámara, 6mm, según las especificaciones mostradas en la Tabla 1. Adicionalmente, se denota cómo la precisión en el cálculo de la distancia se incrementa cuando el brazo robótico que monta la cámara se mueve acercándose a la superficie del objeto que sirve como objetivo. Además, la incapacidad que tiene el sistema para adquirir información de distancia con una convergencia hacia la estabilidad global y no local, radica en la propia configuración del haz de luz de diodos próximos a la lente óptica de la cámara.

4. Ajuste de la Frecuencia de Modulación

4.1. *Análisis de comportamiento de cámara ToF en función de la frecuencia de modulación*

En la Figura 9, se comprueba cómo en ocasiones puede interesar variar la frecuencia de modulación. Frecuencias altas, entorno a 100Mhz, proporcionan errores similares a los 20Mhz al tiempo que permiten garantizar cierta dependencia entre el crecimiento de la frecuencia y el decremento del tiempo de integración. Así, con la misma configuración 'eye-in-hand' que se ha empleado en el experimento anterior, se ha realizado un análisis que permite determinar la frecuencia de modulación más adecuada para estimar la distancia entre cámara y objeto, al tiempo que se fija el parámetro de tiempo de integración. Para llevarlo a cabo, las figuras 10 y 11 representan el comportamiento del error de distancia computado cuando se combinan los parámetros de tiempo de integración y frecuencia de modulación para distintas distancias de trabajo; pero restringiendo el uso de la frecuencia de modulación a los rangos de trabajo mostrados en la figura 9 que indican errores inferiores a 45 mm.

En el experimento de la Figura 8, se realizó un estudio de posibles tiempos de integración para una frecuencia dada, en concreto la preconfigurada por el fabricante, 20Mhz. (Tabla 2). El rango de posibles valores para la frecuencias de modulación que se ha evaluado se sitúa entre [1, 125]Mhz. La idea es ampliar el rango de configuración de la frecuencia de modulación para ver cómo se comporta la cámara ToF en estos casos. Para estimar la distancia al objeto, se ha segmentado el objeto según el método comentado en el apéndice. Éste permite detectar los píxeles que delimitan la región del objeto, determinando la distancia al observador y permitiendo seguir el objeto en la imagen. Los resultados de este experimento se muestran en la Figura 10 y Figura 11. En ambas figuras, se muestra y compara qué relación existe entre el error con respecto a la frecuencia de modulación empleada, el tiempo de integración y la distancia de trabajo. Los valores de error deseables se han marcado en las figuras en color verde y en su defecto en color amarillo. En dichas figuras, únicamente se han representado, las estimaciones cuando el error tolerable ha sido inferior a 45mm. Es decir, cuando es inferior al 5% de la distancia de trabajo, en el experimento que nos ocupa. No obstante, cabe hacer notar que en otros escenarios, distinto al control visual, dónde la distancia de trabajo sea sustancialmente mayor a 1000mm., puede ser interesante tolerar errores ligeramente superiores.

Por un lado, en la Figura 10 se muestra el error en la estimación de la distancia entre cámara y objetivo cuando se trabaja con frecuencias de modulación bajas, entre [5, 35]Mhz., próximas a la preconfigurada por el fabricante. Por debajo de 5Mhz. y por encima de los 25Mhz., los cálculos de la distancia superan el valor de tolerancia permitidos, 5%. Por otro lado, en la Figura 11, se observa el error cuando se trabaja con frecuencias de modulación altas, es decir alejadas del valor predefinido por el fabricante. En estos casos, las frecuencias por debajo de 85Mhz. y por encima de 115Mhz. incrementan la tolerancia de error de distancia permitida.

Así, en este experimento, como se muestra en la figura 10, el mínimo de error posible se mantiene estable, en torno a los 20Mhz. Aunque es aconsejable obtener pequeñas mejoras modificando la frecuencia de modulación en un intervalo de ±2Mhz. En cualquier caso, esas pequeñas variaciones no modifican el error de tolerancia del fabricante (Tabla 1). Con ello se concluye que a bajas frecuencias de modulación, 20Mhz. es un valor muy recomendable, sea cual fuere el tiempo de integración $\tau_d \in [\tau_{min}, \tau_{max}]$ que se configure, siempre que la distancia de trabajo esté en torno al metro o sea inferior a éste. No obstante, en la figura 11, es posible observar que cuando se trabaja con frecuencias de modulación altas, también aparece otro mínimo local para el error. Más aun, este error mínimo a menudo es un mínimo global.

En concreto, en este segundo experimento, el rango de frecuencias de modulación óptimas de uso se sitúa entre 90-110Mhz. Para estos casos, la distancia máxima que se puede estimar, de acuerdo a (3), disminuye. Por ejemplo, a 800mm., los valores mínimos se dan próximos a 94Mhz., a 600mm. a 99Mhz., y a distancias de 400mm entorno a 107Mhz. Por lo tanto, el error varía significativamente cuando la distancia entre cámara y objeto cambia. Y aunque no se puede definir un modelo matemático que demuestre una fuerte dependencia directa entre τ_d, f_{mod} y distancia, si es posible afirmar que a menor distancia interesa incrementar, ligeramente, la frecuencia de modulación deseada cuando se trabaja en un alto rango de frecuencias. De esta manera, cuando se trabaja con frecuencias altas, próximas a 90-110Mhz., que distan bastante de la frecuencia recomendada por el fabricante, 20Mhz., la precisión en la estimación de distancia entre cámara y objeto mejora. Así, comparando las Figuras 10 y 11, se puede observar cómo mientras que en la Figura 10, los valores de color varían significativamente cuando el tiempo de integración es modificado; no ocurre así, en la Figura 11, donde los colores forman una línea de color invariante para una cierta precisión alcanzada y se mantienen sin variación a pesar de que el tiempo de integración se modifique.

Como conclusión, indicar que el error mínimo que se encuentra en las frecuencias altas tiende a un comportamiento cuadrático, donde las frecuencias adyacentes, f_{mod-1} y f_{mod+1}, en torno a la $(f_{mod})^*$ siempre generan un error mayor que ésta (Figura 11).

4.2. Análisis del error

A continuación, se analiza el comportamiento del error cuando se estima la distancia entre cámara y objetivo empleando un tiempo de integración adaptativo calculado online como se ha comentado en la sección 3. En este análisis, se ha considerado también el parámetro frecuencia de modulación.

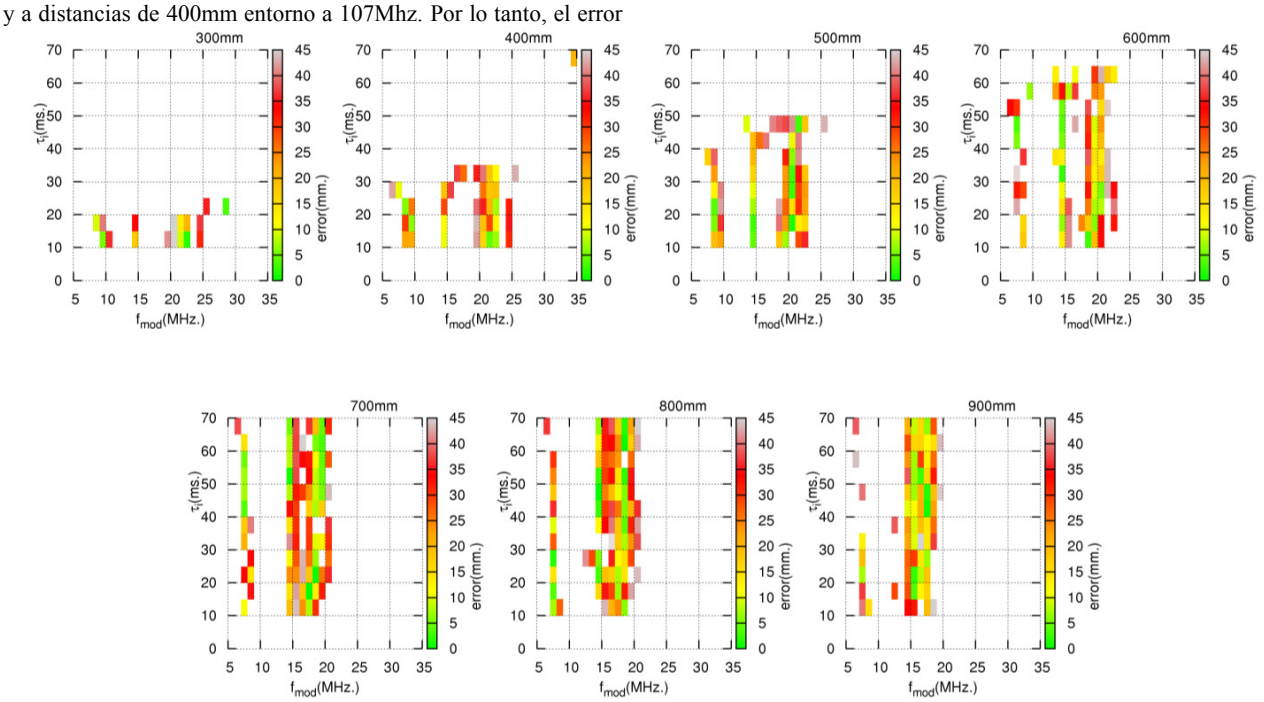

Figura 10: Error en la estimación de distancia entre [300-900]mm. cuando se emplean frecuencias de modulación bajas.

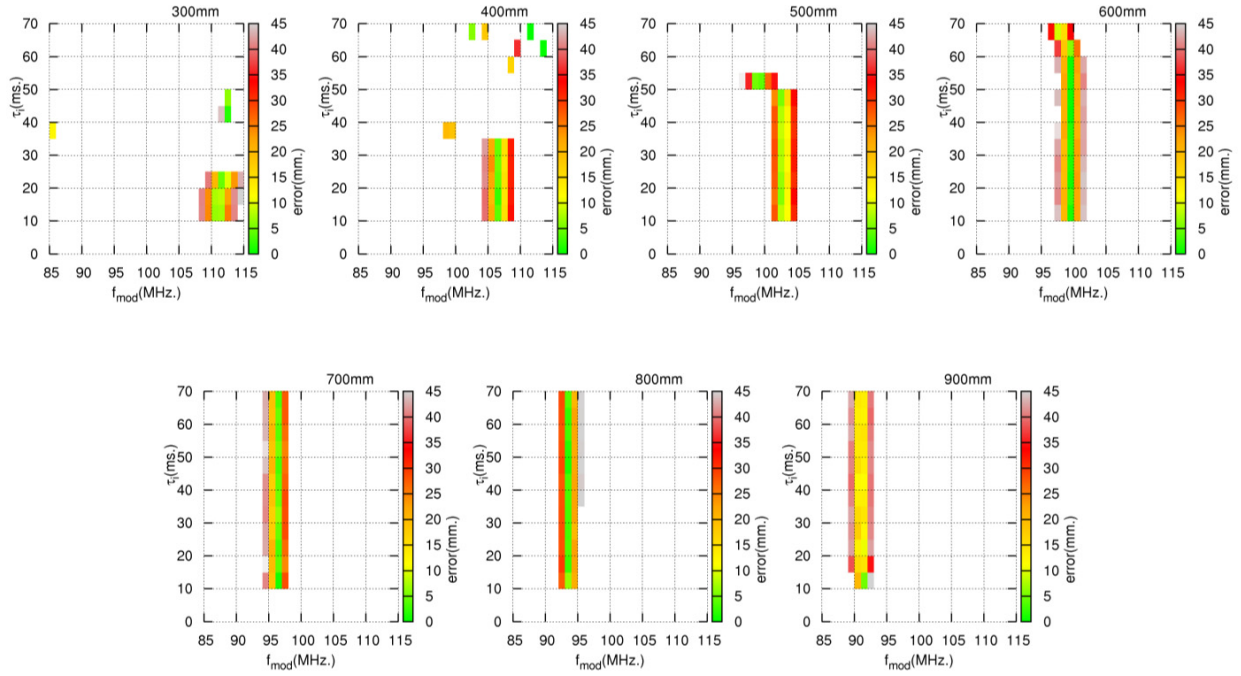

Figura 11: Error en la estimación de distancia [300-900]mm. cuando se emplean frecuencias de modulación altas.

Así, se evalúa el comportamiento del error en ambos casos, cuando se emplean frecuencias de modulación bajas y cuando se emplean frecuencias de modulación altas. En concreto, se ha escogido el valor ideal de los experimentos anteriores (Figuras 10 y 11). Estos valores se sitúan en torno a 17 y 20Mhz. para las bajas frecuencias y en 91 y 99Mhz. para frecuencias altas.

En las Figuras 12 y 13 se muestra la evolución de la distancia calculada cuando la cámara montada en el extremo del robot, sigue a un objeto manteniendo la distancia de trabajo entre cámara y objeto. En el primer caso, Figura 12, la distancia de trabajo se sitúa a 600mm. y en el segundo caso, Figura 13, se sitúa a 900mm. En ambos casos, la diferencia entre la evolución de las medidas de distancia computadas y las reales muestran el error obtenido.

Este error se hace mínimo entre 20 y 30 ms. para el primer caso, tanto si se trabaja a frecuencias bajas, 20Mhz., como frecuencias altas, 99Mhz. El error crece exponencialmente, para valores muy altos del tiempo de integración. En el segundo caso, el error se hace mínimo entre 20 y 30 ms. y aunque se mantiene, se establece a tiempos de integración altos. Además, en ambos casos, la frecuencia de modulación alta presenta una evolución del error más pequeña y estable.

Para finalizar, comentar que se ha estudiado el comportamiento del error como una función Δdistancia. Esta función representa la diferencia sin signo entre el valor real de distancia de trabajo y el estimado. Cada valor de distancia calculado se ha obtenido empleando el tiempo de integración más adecuado, obtenido automáticamente por el algoritmo presentado en la sección 3.2. Del mismo modo, también se ha empleado la frecuencia de modulación subóptima estimada empleando el análisis previo comentado en la sección 4.1. El resultado se muestra en la Figura 14.

En la Figura 14, las medidas calculadas de distancia se han realizado con los valores subóptimos estimados de acuerdo a la distancia de trabajo para el tiempo de integración, τ_k*,y la frecuencia de modulación, (f_{mod})*. El experimento se ha realizado tomando frecuencias de modulación subóptimas, de valor alto, por encima de 63Mhz, y valores de modulación de valor bajo por debajo de 63Mhz y empleando los mismos τ_k* para las mismas distancias en cada caso.

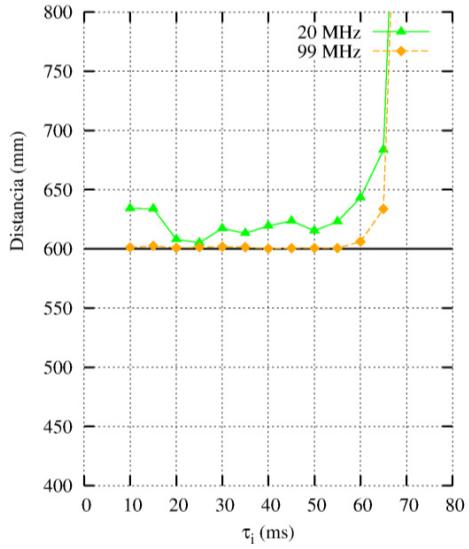

Figura 12: Comparativa en la estimación de distancia I.

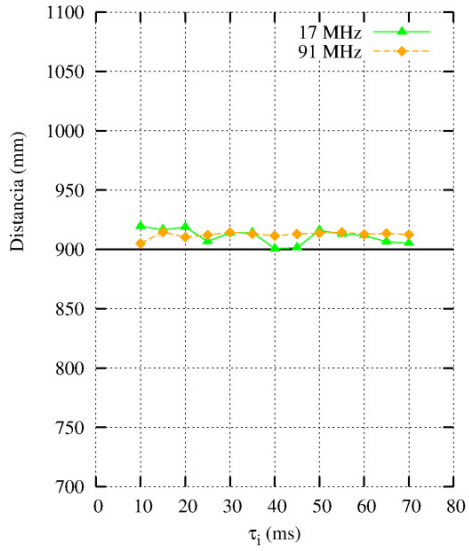

Figura 13: Comparativa en la estimación de distancia II.

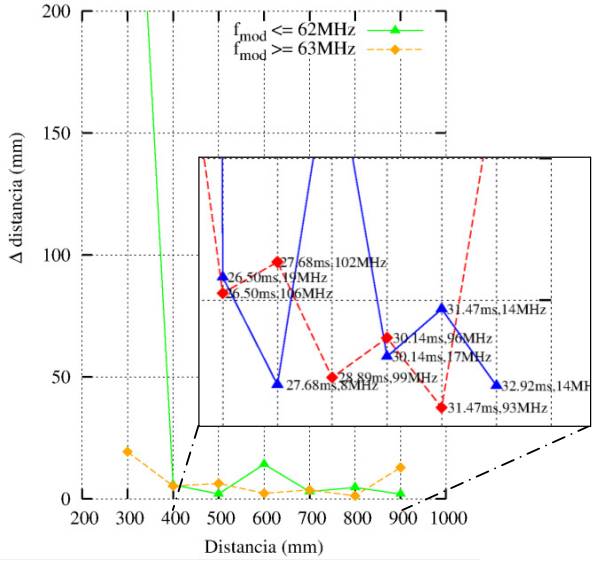

Figura 14: Comparativa en el error obtenido en la estimación de distancia.

5. Conclusión

En este artículo, se ha realizado un análisis exhaustivo y experimental del comportamiento de los parámetros de tiempo de integración y frecuencia de modulación de una cámara ToF. Este análisis, ha permitido demostrar que no sólo el tiempo de integración requiere adaptarse para reducir el error que se produce en la estimación de distancia entre cámara y objetos; sino que también, la frecuencia de modulación se puede adaptar con valores distintos al definido por el fabricante, para mejorar la precisión de esta medida y reducir el error del cálculo de distancia. Como aportación novedosa de este trabajo, por un lado se ha propuesto una metodología general para el autoajuste del tiempo de integración de cámaras ToF. En la actualidad las cámaras ToF requieren, en su mayoría, de un ajuste manual del tiempo de integración en función de la percepción y destreza del operador experto que maneja la cámara. Este ajuste manual obliga a recalibrar la cámara en caso de que varíe la escena. Por otro lado, se ha presentado un rango de valores de alta frecuencia de modulación adecuados para mejorar la precisión de medida respecto a la frecuencia de modulación propuesta por el fabricante. También, se ha demostrado que la frecuencia de modulación no sólo depende de la distancia de trabajo sino también existe una dependencia inversamente proporcional al tiempo de integración a altas frecuencias de trabajo.

Así, a partir de este estudio, se ha presentado una metodología de trabajo para autocalibrar y de este modo estimar valores subóptimos de tiempo de integración y de frecuencia de modulación de cámaras ToF, reduciendo el error y mejorando la precisión en la estimación de distancias. La metodología presentada ha sido probada con éxito en aplicaciones de guiado por control visual basadas en arquitectura ‘eye in hand’. Además, la metodología expuesta es extrapolable a otras aplicaciones de guiado de robots en las que el sistema sensorial se compone de sensores ToF.

English Summary

ToF Camera calibration: an automatic setting of its integration time and an experimental analysis of its modulation frequency

Abstract

The depth perception is essential in many manipulation tasks, visual inspection and robot navigation. The cameras of Time of Flight (TOF) generate range images which provide depth measurements in real time. However, the distance parameter computed from these cameras is strongly dependent on the integration time set for the sensor and the frequency of modulation used by the integrated lighting system. In this paper, a methodology for automatic setting of integration time and an experimental analysis of ToF camera behavior adjusting its modulation frequency is presented. This method has been successfully tested on visual servoing algorithms with architecture 'eye-in-hand' in which the sensory system consists of a ToF camera, in addition this methodology can be applied to other workspaces and scenarios.

Keywords:

Time of flight, calibration, range image, robotic perception, 3d-cameras.

Agradecimientos

Este trabajo ha sido co-financiado por el Gobierno regional de la Generalitat Valenciana, Universidad de Alicante y CICYT través de los proyectos GV2012/102, GRE10-16 y DPI2012-32390.

Referencias

Bouguet, J. Y., 2000. Pyramidal implementation of affine Lucas Kanade feature tracker. Intel Corporation- Microprocessor Research Labs, OpenCV Library.

Chaumette, F., Hutchinson, S., 2006. Visual servo control. I. Basic approaches. IEEE Robotics and Automation Magazine 13, IEEE Press, pp. 82-90.

Chiabrando, F., Chiabrando, R., Piatti, D., Rianudo, F., 2009. Sensors for 3d imaging: metric evualuation an calibration of CCD/CMOS time-of-flight camera. Sensors 9(9), pp. 10080-10096. DOI: 10.3390/s91210080

Distante, C., Diraco, G., Leone, A., 2010. Active range imaging dataset for indoor surveillance. In Proc. of British Machine Vision Conference (BMVC), BMVA Press, vol. 2, pp.1-16.

Foix, S., Aleny, G., Torras, C., 2011. Lock-in time-of-flight (ToF) cameras: a survey. IEEE Sensors Journal 11(3), pp. 1917-1926. DOI: 10.1109/JSEN.2010.2101060

Frank, M., Plaue, M., Rapp, H., Köthe, U., Jähne, B., Hamprecht, F.A., 2009. Theoretical and experimental error analysis of continous-wave time-of-flight range cameras. Optical Engineering 48(1), pp. 13602-13618.

Fuchs, S. Hirzinger, G., 2008. Extrinsic and depth calibration of ToF-cameras. In Proc. of Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1-6. DOI: 10.1109/CVPR.2008.4587828

Garcia, F., Aouada D., Mirbach, B., Solignac T., Ottersten B., 2011. Real-time hybrod ToF multi-camera rig fusion system for depth map enhancement. In Proc. of Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1-8. DOI: 10.1109/CVPRW.2011.5981740

Gil, P., Pomares, J., Torres, F., 2010. Analysis and adaptation of integration time in PMD camera for visual servoing. In Proc. of 20th International Conference on Pattern Recognition (ICPR), IEEE Press Society, pp. 311-315. DOI: 10.1109/ICPR.2010.85

Herrera, D.C., Kannala, J., Heikkila, J., 2011. Accurate and practical calibration of a depth and color camera pair. In Proc. of 14th International Conference on Computer Analysis of Images and Patterns (CAIP), vol 2, Ed. Springer-Verlag Berlín, Heidelberg, pp. 437-445.

Hussman, S., Liepert, T., 2009. Three-dimensional tof robot vision system. IEEE Transactions on Instrumentation and Measurement 58(1), pp. 141-146. DOI: 10.1109/TIM.2008.928409

Hussman, S., Edeler, T., 2010. Robot vision using 3d tof systems. En: Ales Ude (Ed.), Robot Vision. Intech Press, pp. 293-306.

Kakiuchi, Y., Ueda, R., Kobayashi, K., Okada, K., Inaba, M., 2010. Working with movable obstacles using on-line environmet perception reconstruction using active sensing and color range sensor. In Proc. of International Conference on Intelligent Robots and Systems (IROS), IEEE Press, pp. 1696-1701. DOI: 10.1109/IROS.2010.5650206

Kim, Y.M., Chan, D., Theobalt, C., Thrun, S., 2008. Design and calibration of a multi-view ToF sensor fusion system. In Proc. of 22nd Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1524-1530. DOI: 10.1109/CVPRW.2008.4563160

Kisler, T., Gil, P., 2011. Detección y seguimiento de objetos sólidos con cámaras ToF. Actas de XXXII Jornadas de Automática (JA), CEA-IFAC Actas. Sevilla (Spain).

Khoshelham, K. 2011. Accuracy analysis of Kinect depth data. En: D.D. Lichti and A.F. Habbib (Ed.). In Proc of ISPRS Journal of Photogrammetry and Remote Sensing-Workshop on Laser Scanning, vol. 38(5), pp. 29-31.

Kolb, A., Barth, E., Koch, E., Larse, R., 2010. Time-of-flight Cameras in Computer Graphics. Computer Graphics Forum, vol. 29(1), pp. 141-159. DOI: 10.1111/j.1467-8659.2009.01583.x

Kuehnle, J.U., Xue, Z., Sotz, M., Zoellner, J.M., Verl, A., Dillmann, R., 2008. Grasping in depth maps of time-of-flight cameras. In Proc. of IEEE International Workshop on Robotic and Sensors Environments (ROSE). pp. 132-137. DOI: 10.1109/ROSE.2008.466914

Lai, K., Liefeng Bo, Xiaofrng Ren, Fox, D., 2011. Spares distance learning for object recognition combining RGB and depth information. In Proc. of International Conference on Robotics and Automation (ICRA), IEEE Press Society, pp. 4007-4013. DOI: 10.1109/ICRA.2011.5980377

Lichti, D., 2008. Self-calibration of a 3D range camera. In Proc of International Society for Photogrammetry and Remote Sensing 37(3), pp.1-6

Lichti, D. Rouzaud, D., 2009. Surface-dependent 3d range camera self-calibration. En: A Beraldin, G.S. Cheok, M. McCarthy (Ed.), In Proc. of SPIE vol. 72390, pp. . DOI: 10.1117/12.805509

Lichti, D., Kim, C., 2011. A comparison of three geometric self-calibration methods for range cameras. Remote Sensing 11(3), pp. 1014-1028. DOI: 10.3390/rs3051014

Lindner, M., Kolb, A., Ringbeck, T., 2008. New insights into the calibration of ToF-sensors. In Proc. of 22nd Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1603-1607. DOI: 10.1109/CVPRW.2008.4563172

Lindner, M., Schiller, I., Kolb, A., Koch, R., 2010. Time-of-flight sensor calibration for accurate range sensing. Computer Vision and Image Understanding vol. 114(2), pp. 1318-1328. DOI: 10.1016/J.CVIU.2009.11.002

May, S., Werner, B., Surmann, H., Pervölz, K., 2006. "3d time-of-flight cameras for mobile robotics. In Proc. of International Conference on Intelligent Robots and Systems (IROS), IEEE Press, 790-795, DOI: 10.1109/IROS.2006.281670

May, S., Fuchs, S., Droeschel, D. Holz, D., Nüchter, A., 2009. Robust 3d-mapping with time-of-flight cameras. In Proc. of International Conference on Intelligent Robots and Systems (IROS), IEEE Press Society, pp 1673-1678.

May, S., Droeschel, D., Holz, D., Fuchs, S., Malis, E., Nüchter, A., Hertzberg, J., 2009. Three-dimensional mapping with time-of-light cameras. Journal of Field Robotics. Special Issue on Three-dimensional Mapping Part 2, 26(11-12), pp. 934-965. DOI: 10.1002/ROB.20321

Mufti, F., Mahony, R., 2011. Statistical analysis of signal measurement in time of flight cameras. Journal of Photogrammetry and Remote Sensing (ISPR) vol. 66(5), pp. 720-731. DOI: 10.1016/J.ISPRSJPR.2011.06.004

Pattison, T., 2010. Quantification and description of distance measurement errors of a time-of-flight camera. M.Sc. Thesis. University of Stuttgart, Stuttgart (Germany).

Pomares, J., Gil, P., Torres, F., 2010. Visual control of robots using range images. Sensors 10(8), pp. 7303-7322. DOI: 10.3290/s100807303

Rapp, H., Frank, M., Hamprecht, F.A., Jähne, B., 2008. A theoretical and experimental investigation of the systematic errors and statistical uncertainties of time-of-flight-cameras. International Journal of Intelligent Systems Technologies and Applications vol. 5(3-4), pp. 402-413. DOI: 10.1504/IJISTA.2008.021303

Shahbazi, M., Hoimayouni, S., Saadatseresht, M., Sattari, M., 2011. Range camera self-calibration based on integrated bundle adjustment via joint setup with a 2d digital camera. Sensors 11(11), pp. 8721-8740. DOI: 10.3290/s110908721

Schaller, C., 2011, Time-of-Flight-A new Modality for Radiotherapy, M.Sc. Thesis. University Erlangen-Nuremberg, Erlagen (Germany).

Schiller, I., Beder, C., Koch, R., 2008. Calibration of a PMD-camera using a planar calibration pattern together with a multi-camera setup. In Proc. of ISPRS Journal of Photogrammetry and Remote Sensing vol. 37, pp. 297-302.

Schwarz, L., Mateus, D., Castaneda, V., Navab, N., 2010. Manifold learning for ToF-based human body tracking and activity recognition. In Proc. of British Machine Vision Conference (BMVC), BMVA Press, pp.1-11. DOI: 10.5244/C.24.80

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A., 2011. Real-time human pose recognition in parts from single depth images. In Proc of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1297-1304.

Smisek, J., 2011. *3D Camera Calibration*. MSc. Thesis. Czech Technnical Univesity, Prague (Czech).

Weyer, C.A., Bae, K.H., Lim, K., Lichti, D., 2008. Extensive metric performance evaluation of a 3D range camera. In Proc. of ISPRS Journal of Photogrammetry and Remote Sensing vol.37(5), pp.939-944.

Wiedemann M., Sauer M., Driewer F. Schilling K., 2008. Analysis and characterization of the PMD camera for aplication in mobile robots. M. J. Chung and P. Misra (Ed.). In Proc. of 17th World Congress of the International Federation of Automotic Control, IFAC Press, pp.13689-13694.

Zhang, Z., 2000. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), pp. 1330-1334. DOI: 10.1109/34.888718

Zhu, J., Wang, L., Yang, R., Davis, J., 2008. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In Proc. of Computer Vision and Pattern Recognition (CVPR), IEEE Press Society, pp. 1-8. DOI: 10.1109/CVPR.2008.4587761

Zhu, J., Yang, R., Xiang, X., 2011. Eye contact in video conference via fusion of time-of-flight depth sensor and stereo. 3D Research 2(3), pp. 1-10. DOI: 10.1007/3Dres.03(2011)5

Zinber, T., Schmidt, J., Niemann, H., 2003. A refined ICP algorithm for robust 3d correspondence estimation. In Proc. of Conference on Image Processing (ICIP), IEEE Press, pp. 695-698.

Apéndice A. Segmentación y detección de objetos en imagen de rango ToF.

Para la calibración de los parámetros de tiempo de integración y frecuencia de modulación, es necesario detectar los píxeles que pertenecen a objetos en la imagen. Esto se ha llevado a cabo mediante un proceso de segmentación por 'blobs' de distancia. Éste se realiza estimando la distancia media de los píxeles de la imagen según (3), donde la distancia computada, d_i, es la distancia para cada i-píxel que pertenece a la imagen. Entonces, si la distancia de una determinada región de interés en la imagen es menor que la distancia media de la imagen (cerca de la cámara) se considera que esta región está más próxima a la cámara que al fondo y se etiquetan todos sus píxeles como pertenecientes a un objeto. Aunque la distancia media no representa la distancia real a la que se encuentra el fondo, este tipo de estimación permite segmentar objetos de fondo sin tener información a priori sobre la presencia de número y tamaño de objetos y posición del fondo.

Por otro lado, la distancia computada, d_i, para cada i-píxel de coordenadas (u,v), calculadas en (3) representa una distancia radial pero no la profundidad real entre cámara y objetivo. Por este motivo, estos parámetros (u,v,d) deben ser transformados en coordenadas cartesianas (x,y,z) que representen la posición relativa de cada píxel de la imagen al centro óptico del sensor. Para realizar esta transformación se asume una adaptación del modelo de proyección de pinhole, considerando que la longitud focal (f_u,f_v) está escalada en las direcciones (u,v) y que las coordenadas del punto principal son (u_0,v_0). Estos valores se han obtenido mediante un proceso de calibración de los parámetros intrínsecos del sensor y la óptica (Zhang *et al.*, 2000). En este trabajo, sin considerar los factores de distorsión de la lente y adaptando el modelo de pinhole, se puede obtener:

$$z_i = d_i \cdot \frac{1}{\sqrt{1+\frac{(u-u0)^2}{f_u{}^2}+\frac{(v-v0)^2}{f_v{}^2}}}$$

El esquema general del algoritmo implementado para la detección y segmentación de objetos se puede resumir como se indica a continuación siguiendo la notación de la tabla 3.

Algoritmo 3: Detección y segmentación con cámara ToF

Para k=1 hasta N ∀ N posiciones del robot

i. Mover el robot a la posición pos_{des}
ii. Adquirir mapa de profundidades d_i ∀ pixel-i ϵ I
iii. Adquirir umbral de distancia media t_{dist}
iv. Distribuir los n puntos semilla P_{sem} de forma equidistante en la imagen I
v. ∀ pixel-i ϵ P_{sem}
   a. Si pixel-i ϵ P_{del}: { pixel-i | i ϵ o, o ϵ O}: borrar(pixel-i)
   b. sino Si $d_i < t_{dist}$
      i. ∀ p_{vec} añadir P_{vec} a o si $d(p_{vec}) < u_{dist}$
      ii. Si ∃ pixel-i ϵ P_{vec_merge}: { pixel-i | $d_i < u_{dist}$ ∧ pixel-i ϵ o, o ϵ O}: borrar objetos
   c. sino borrar(p)
vi. borrar(o), ∀ o ϵ O_{del}: {o | region(o) < u_{region}}
vii. Calcular $c_{gravedad}$ ∀ o ϵ O
viii. Usar Lukas-Kanade para seguir objetos, ∀ o ϵ O

Para reducir el proceso de segmentación y cálculo de 'blobs' la imagen se muestra a 1/10 puntos distribuidos regularmente por toda la imagen. Posteriormente, se aplica un algoritmo de inundación-relleno ('flood-fill', en la literatura anglosajona) basado en crecimiento de regiones y relleno para detectar la presencia de objetos, *o*. Para estabilizar la detección y discernir qué pertenece al fondo (mesa) y qué es objeto, se han tomado máscaras para evaluar el entorno de vecindad de conectividad-4.
De este modo, se evalúan 5 píxeles por cada muestra. Las muestras evaluadas cuyos píxeles, *pixel-i*, no pertenecen a un objeto detectado se eliminan, P_{del}, y de esta manera, no tienen que volver a ser calculadas de nuevo.

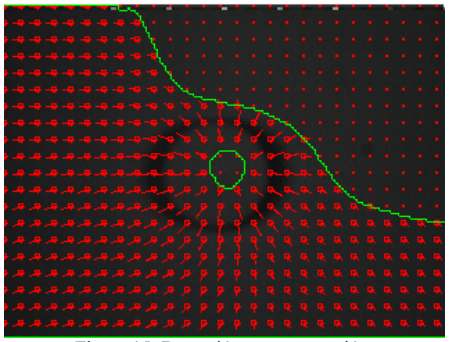

Figura 15: Detección y segmentación.

Después, de detectar y localizar los objetos, *O*, presentes en la imagen inicial, iteración k=0 (posición inicial del robot-cámara), en una segunda etapa, se requiere encontrar esos mismos objetos en el resto de iteraciones k>1. Cada iteración k, representa el instante de tiempo en el que se adquiere una imagen. Además, durante la secuencia de tiempo en la que se adquieren imágenes, el observador (robot-cámara) o uno o varios objetos representados en la imagen pueden haberse movido, provocando desplazamientos en la imagen.

Para detectar la presencia y localización de los objetos en el resto de iteraciones, sin necesidad de calcular de nuevo, el algoritmo de segmentación comentado anteriormente, se emplea la técnica de seguimiento por flujo óptico de Lucas-Kanade (Bouguet, 2000), cuyo resultado se observa en la figura 15.

Tabla 3: Notaciones algoritmo 3

I:	Imagen; i: Píxel de la imagen
$p_{u,v}$:	Valor del píxel en la posición (u,v)
O:	Lista de objetos; o: Objeto
O_{del}:	Objectos a borrar; P_{del}: Puntos a borrar
P_{vec}:	Conjunto de puntos vecinos conectados mediante conectividad-4
pos_{des}:	Posición del objetivo
P_{sem}:	Puntos de comienzo para el proceso de segmentación
d():	distancia Euclidea; u_{dist}: Umbral de distancia